Implementing a Condor pool using a Green-IT policy

D. Gubb, V. Holmes, I. Kureshi, S. Liang, Y. James
University of Huddersfield, Queensgate, Huddersfield HD1 3DH, UK

ABSTRACT

1. Background

High Throughput Computing (HTC) systems are designed to utilise available resources on a network of idle machines in an institution or organization by cycle stealing. It provides an additional 'free' resource from the existing computing and networking infrastructure for modelling and simulation requiring a large number of small jobs, such as applications from biology, chemistry, physics, and digital signal processing. At the University of Huddersfield, there are thousands of idle laboratory machines that could be used to run serial/parallel jobs by cycle stealing. Our HTC system, implemented in Condor [1], is part of the Queensgate Campus Grid (QGG) [2] that consists of a number of dedicated departmental and university computer clusters.

Condor is an excellent HTC tool that excels in cycle stealing and job scheduling on idle machines. However, only idle powered machines can be used from a networked pool. Many organizations deploy power saving mechanisms to try to reduce energy consumption in their systems, and power down idle resources, using rigid and inflexible power management policies. The University of Huddersfield Computing Services use the Energy Star EZ GPO power saving tool that runs as a Windows service and detects how long the computer has been idle. Then it allows the computer to first turn off the screen and then go into hibernation.

Our research and development work is focused on implementing a HTC system using Condor to work within a "green IT" policy of a higher education institutions that conform to green IT challenges for a multiplatform, multidiscipline user/ resource base. This system will allow Condor to turn on machines that may have gone to sleep due to lack of usage when there is a large queue of pending jobs. The decision to utilise dormant resources will be made on a variety of factors such as job priority, job requirements, user priority, time of day, flocking options, queue conditions etc. Good practice scheduling policies would need to be devised that would work within this "green IT" pool.

Condor is a cross-platform tool that can run on Windows, Linux and Macintosh operating systems. Most of the University of Huddersfield computer laboratories are running Windows 7. Some of the serial applications that currently run on the QGG Eridani cluster have been ported into Windows and run successfully. However, a number of applications still run natively in Linux, supporting the multi-platform, multi-discipline research community. By using Condor we are making more efficient use of the current computing resources distributed across the University of Huddersfield to support our growing research community.

Many HE institutions use Condor as part of their HPC/HTC provision. We have conducted a survey to determine its use and identify good practices that can be deployed at the University of Huddersfield.

2. Survey Results

A survey was commission by the Campus Grids Special Interest group (CG-SIG), to determine the effectiveness of having and maintaining a Condor pool within a research institution. The survey was then advertised and filled in by members of the CG-SIG and members of the Northwest Grid.

The purpose of the survey was to identify the following:

- Number of cores in a Condor pool

- Operating system used and implementation
- Condor pool administration
- Payment scheme

The result of the survey has demonstrated that most of the Condor pools have 1000 to 2000 cores and some greater than 2000 cores. The large numbers of cores are result of the fact that most modern desktop computers deployed in laboratories having at least 4 cores.

Just over half of the Condor pools in the survey ran the Linux client for the Condor execute machine in the teaching laboratories. This is achieved by either flocking to different pools setup by researchers or by dual booting the Windows lab machine into a flavour of Linux and then running the Linux client. The other institutions ran the Windows Condor client because it is the easiest way to deploy it across large organisations.

Most organisations use a 100Mbps connection and some had 1Gbps interconnects. The reason for a predominantly 100Mbs network is that typically laboratory machines do not usually require a high level of performance. Hence it is not a valid option to upgrade it to 1 Gbps because of the extra cost involved.

The number of researchers per organisation, who actively use the Condor pool, range between 10 and 120 active users with an average of 45 active users per organisation. Condor pools are not generally used for teaching or assignment except for a few cases of final year projects.

Each Condor pool offers a support site (or Wiki page) providing help and advice to the users of the Condor pool. Over half of the organisation offered training sessions for Condor with a quarter offering code support session to try and troubleshoot any bugs in programs, trying to ensure that the users are getting the best possible use of the Condor pool.

Some of the institutions are proactive in attracting new users to Condor pools by advertising them on the institutions' Condor website and encouraging current users to share their experience. A quarter of the responders never make any effort to effectively advertise the Condor pool. One organisation produces a newsletter every month to let the users know of any upgrade and updates in the pool.

Condor does not have any effective administration or management tools so administrators tend to create their own scripts and programs to monitor the status of the Condor pool.

Typical methods of providing security on a Condor pool are restricting which machines can submit jobs to the Condor pool, or using SSH keys to restrict access with use of an LDAP authentication.

Only half of the respondents use power management with their Condor pools. One of the institutions just uses the standard Condor settings which put the machines into hibernation after 15 minutes of being idle. Another institution tried to implement the Condor power management mechanism that was not very successful. Currently they use their own home-grown solution so that they can turn machines off and then power them up depending on the usage.

Almost all of the responses believed that Condor is a valuable, low cost resource that can help to support research. Out of all the responses in the survey there has only been one negative comment regarding the memory-intensive jobs. It was noticed that such jobs ran slower when a laboratory machine was used with Condor running in the background. None of the other users noticed that Condor was running in the background due to Condor's effective cycle-stealing policies.

Also Condor has provided an alternative resource that can be used to alleviate the number of jobs that have to go through the High Performance Computer Clusters as well as being the best alternative to run embarrassingly parallel tasks.

Normally there is very little financial support provided by the institutions to invest into the improvement of Condor.

One important point raised from one of the submissions is that there is a limited amount of support for the Windows Condor client as compared to the Linux Condor client. Taking into account that university computer laboratories tend to run Windows as their main operating system, this is even more important to address.

3. Huddersfield Deployment

At the University of Huddersfield, the main use of the Condor pool is to provide an extra computing resource alongside the existing computing clusters on the university campus.

The Condor Master and Job Scheduler are located on a Linux machine with a dual socket Opteron 2350 with 16Gbytes of RAM. The majority of the computers have at least an Intel quad core processor or better and are running Windows 7.

We have devised a script using a modified version of the solution at the University of Liverpool [3].This script enables load balancing in the Condor pool allowing a dual boot of machines using the Windows Condor Client, and virtualising Linux using Pools of Virtual Boxes (PoVB). The script will wake up the machines using Wake in LAN (WoL). It runs every 15 minutes as a Cron job on the Condor Master node (see Figure 1)

```
Number of Busy Machines = 151
Number of idle jobs = 783
Number of idle machines = 11
Number of idle Win7 = 10
Number of idle WinXP = 0
Number of require machines = 772

Number of idle Windows jobs = 783
Number of idle Windows machines = 10
Number of Windows Machines required = 773

Number of idle Linux jobs = 0
Number of idle Linux Machines = 1
Number of Linux Machines Reqired = -1
--------------------------------------------------

##################################################
#                                                #
# Number of required machines = 772 #
# Machines have been woken                       #
#                                                #
##################################################
```

Figure 1 Huddersfield Power Management Script

A set of rules were implemented in the script to support ‘green IT’ policy. When a lab machine is turned on, the Condor Windows service runs in the background waiting until the computer has been idle for 5 minutes and then run Windows Condor jobs. If there are more than 12 idle cores then Condor will remotely turn off the Windows Condor client with Microsoft Remote Procedure Call and then turn on the PoVB service. PoVB continues to turn on the virtual machine until either there are no more machines idle or there are no more Linux jobs. If there are too many PoVB Clients idle and not enough Windows clients, then the PoVB service is turned off and the Windows service is started.

The advantage of using the Condor Windows client is that it can directly use the hardware of the Windows machines, which is important given that we do not have a significant Windows HPC platform. This approach will make up the shortfall in Windows HPC in our campus grid QGG.

Also, we have identified that it was essential to have Linux environment within the Condor pool because the majority of the users of the QGG use the Linux clusters, both on and off the campus. The latest version of PoVB is a 64-bit version that allows the users to use as much of the machines’ performance as possible. This is significant because less of the machines’ true potential will be used with a virtual machine as opposed to an actual Linux machine.

In addition, the Windows Condor service and the PoVB have been configured to have an extra job slot, so that when the entire machine is free the user can request to use it all. This allows the users to utilise the entire RAM of the machine for more memory intensive jobs. They can also use the network mounted MPI libraries to run parallel applications. This is particularly important because, as the users begin to use the QGG, their calculations do not normally require more than 4 cores. As technology moves on and more cores are becoming the standard in desktop machines, the ability to perform parallel applications will increase. This allows the smaller simulations to run on the Condor pool, and frees up the computer clusters for users with larger simulations.

4. Summary

The advantage of having a Condor Pool that can conform to a green IT policy, perform small parallel applications and provide dual operating systems with a low cost of ownership opens up the opportunity to have a HTC/HPC environment in HE institutions. This will encourage organisations to consider how their users are using existing computer clusters and HPC resources, whether they actually need to have small-dedicated clusters, or it is feasible to use Condor pools that could save money in power and infrastructure.

As a result of our work and deployment of our existing Condor pools that are currently used by our researchers, a decision was made to deploy Condor on all computers in the Queensgate campus at the University of Huddersfield from September 2012.

5. References

[1] "Condor High Throughput Computing," July 2012. [Online]. Available: http://research.cs.wisc.edu/condor/. [Accessed 26 July 2012].

[2] V. Holmes and I. Kureshi, "Huddersfield University Campus Grid: QGG of OSCAR Clusters," Journal of Physics: Conference Series, vol. 256, pp. 012 - 022, 2010.

[3] I. C. Smith, 2010 August 2010. [Online]. Available: http://www.liv.ac.uk/csd/escience/condor/power_save.pdf. [Accessed 2011 October 12].